

Figure 1

0940189:082701

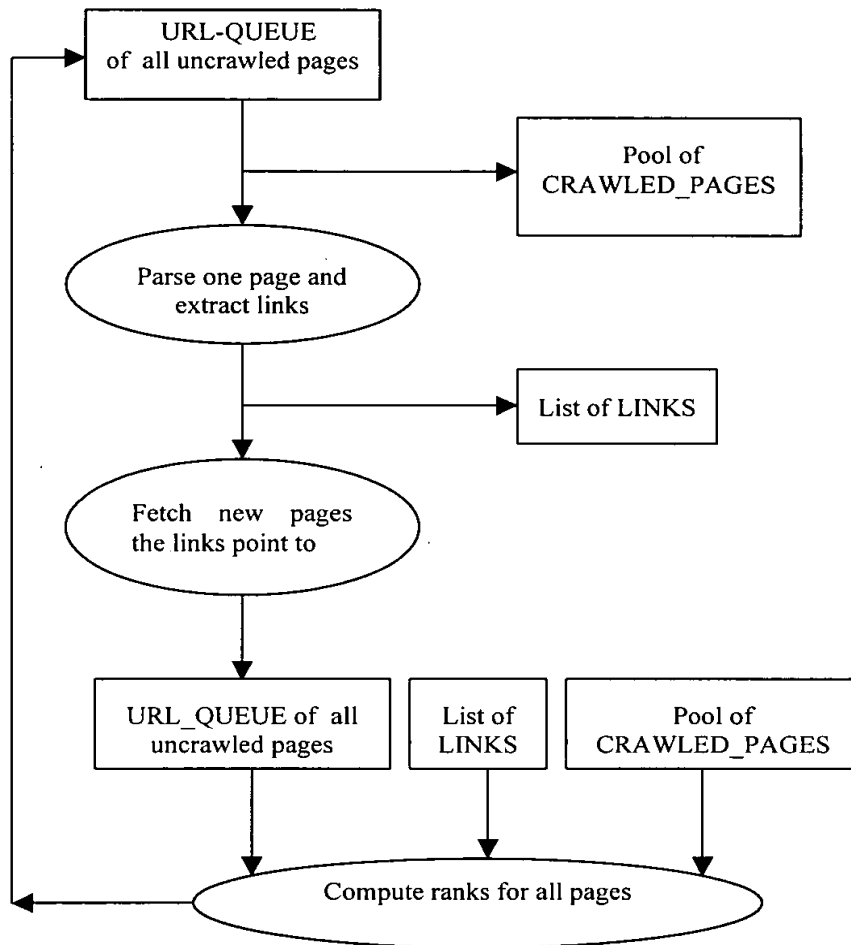


Figure 2

```

Enqueue(URL_QUEUE, starting_url);

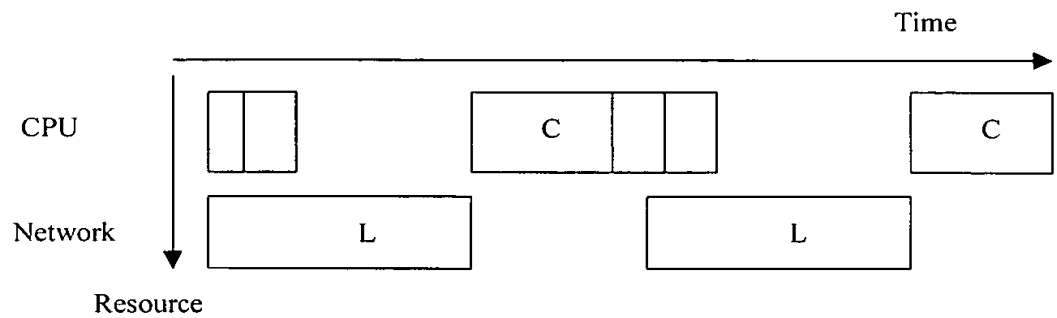
While (! Empty(URL_QUEUE)) {
    url = Dequeue(URL_QUEUE);
    page = Crawl_page(url);
    Enqueue(CRAWLED_PAGES, (url, page));
    PAGE_TO_FETCH_LIST = Extract_urls(page);
    For each u in PAGE_TO_FETCH_LIST {
        Enqueue(LINKS, (url, u));
        If ((u not in URL_QUEUE) and ((u, -) not in CRAWLED_PAGES)) {
            Fetch_page(u);
            Enqueue(URL_QUEUE, u);
        }
    }
    Compute_PageRank(URL_QUEUE+CRAWLED_PAGES);
    Reorder_queue(URL_QUEUE);
}

```

Figure 3

104280" 88101660

0940188-082701
T02280"88T04560



- F: Initiating page fetches (protocol management)
- R: Receiving/taking pages from incoming queue
- C: Computation
- L: Loading pages from remote servers through network

Figure 4

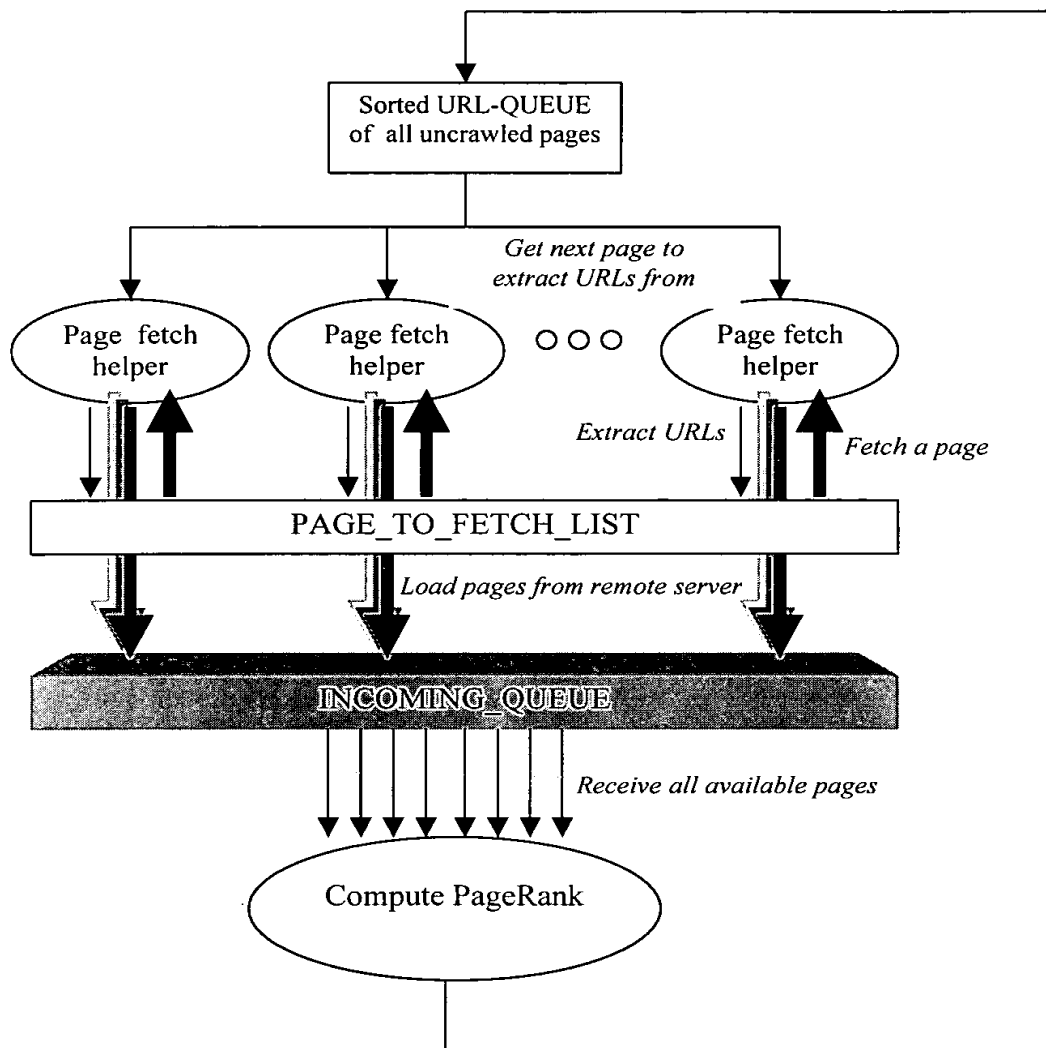


Figure 5